SOLUTIONS TO SELECTED ADDITIONAL EXERCISES FOR MATHEMATICS 132 — Part 4

Winter 2004

IV. Complex linear algebra

IV.2 : Complex matrices

1. Let χ be the map sending a vector

$$\mathbf{z} \;=\; (z_1, \; \cdots \;, z_n\,) \;\in\; \mathbb{C}^n$$

to the vector with conjugate entries:

$$\chi(\mathbf{z}) \;=\; (\overline{z_1}, \; \cdots \;, \overline{z_n}\,)$$

Give examples of subspaces W of the complex vector space $\mathbb{C}^n$ such that $\chi(W) = W$ and $\chi(W) \neq W$.

SOLUTION.

The spans of the various subsets of unit vectors are mapped to themselves under χ. On the other hand, consider the subspace of $\mathbb{C}^2$ spanned by the vector $(1, i)$. We claim that this subspace is not sent to itself under χ. It suffices to show that $\chi(1, i) = (1, -i)$ is not a scalar multiple of $(1, i)$. But suppose that $z\,(1, i) = (1, -i)$. The left hand side is equal to $(z, \, i\,z)$, so the vector equation reduces to the two scalar equations $z = 1$ and $z\,i = -i$. Since the second of these equations implies that z must be equal to -1, it follows that there is no common solution to the two scalar equations and therefore no multiple of the vector $(1, i)$ is equal to $(1, -i)$. This means that the subspace spanned by the first vector is not sent to itself under the conjugation map.∎

2. Let $\mathcal{A} = \{\,\mathbf{v}_1, \; \cdots \;, \mathbf{v}_k\,\}$ be a set of vectors in $\mathbb{C}^n$ with **real** coordinates. Prove that $\mathcal{A}$ is linearly independent in $\mathbb{C}^n$ if and only if it is linearly independent when viewed as a subset of $\mathbf{R}^n$. [*Hint:* Why is every vector in $\mathbb{C}^n$ uniquely expressible in the form $\mathbf{x} + i\,\mathbf{y}$ where both $\mathbf{x}$ and $\mathbf{y}$ have real coordinates?]

SOLUTION.

By construction, every vector in $\mathbb{C}^n$ has the form $\mathbf{x} + i\,\mathbf{y}$ for some vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbf{R}^n$, and two vectors $\mathbf{x} + i\,\mathbf{y}$ and $\mathbf{u} + i\,\mathbf{v}$, where $\mathbf{x}$, $\mathbf{y}$, $\mathbf{u}$, $\mathbf{v} \in \mathbf{R}^n$, are equal if and only if $\mathbf{x} = \mathbf{y}$ and $\mathbf{u} = \mathbf{v}$.

Suppose that the given vectors are linearly independent in $\mathbb{C}^n$. This means that the only way of writing $\mathbf{0}$ as a linear combination $\sum_j c_j\,\mathbf{v}_j$ with $c_j \in \mathbb{C}$ is the trivial linear combination where each $c_j = 0$. Since every real linear combination is a complex linear combination, it follows that the only way of writing $\mathbf{0}$ as a real linear combination is the trivial one, and hence the vectors are also linearly independent over the reals.

Conversely, suppose that the given vectors are linearly independent in $\mathbf{R}^n$, and suppose we have a complex linear combination $\sum_j c_j\,\mathbf{v}_j$ that is equal to $\mathbf{0}$. If for each j we write $c_j = a_j + i\,b_j$ where each a_j and b_j is real, then we may rewrite this as follows:

$$\mathbf{0} + i\,\mathbf{0} \quad = \quad \sum_j c_j\,\mathbf{v}_j \quad = \quad \sum_j (a_j + i\,b_j)\,\mathbf{v}_j \quad =$$

$$\left(\sum_j a_j\,\mathbf{v}_j\right) \quad + \quad i \cdot \left(\sum_j b_j\,\mathbf{v}_j\right)$$

Equating the real and imaginary parts of the left and right hand sides we obtain the equations

$$\mathbf{0} \quad = \quad \sum_j a_j\,\mathbf{v}_j \quad = \quad \sum_j b_j\,\mathbf{v}_j$$

where the coefficients a_j and b_j are all real. Since the set of vectors $\mathcal{A}$ is linearly independent over the real numbers, this means that $a_j = 0$ for all j and $b_j = 0$, which in turn implies that

$$c_j \quad = \quad a_j + i\,b_j \quad = \quad 0 + i\,0 \quad = \quad 0$$

for all j; *i.e.*, the set $\mathcal{A}$ is linearly independent over $\mathbb{C}$.■

3. If V is a complex vector space, then one can also view V as a vector space over the real numbers by defining the real scalar product of $c \in \mathbf{R}$ and $\mathbf{v} \in V$ to be the complex scalar product $(c + 0\,i)\mathbf{v}$. Suppose that $W \subset \mathbb{C}^n$ is a complex subspace and $\mathcal{A} = \{\,\mathbf{v}_1,\ \cdots\ ,\mathbf{v}_k\,\}$ is a basis for W as a complex vector space. Prove that $\mathcal{A} \cup i \cdot \mathcal{A}$ is a basis for W as a real vector space, where

$$i \cdot \mathcal{A} \quad = \quad \{\,i\,\mathbf{v}_1,\ \cdots\ , i\,\mathbf{v}_k\,\}\ .$$

What does this say about the relationship between the real and complex dimensions of W?

SOLUTION.

If $\mathbf{w} \in W$, then $\mathbf{w} = \sum_j\ c_j\,\mathbf{v}_j$ for suitable complex numbers c_j. If $c_j = a_j + b_j\,i$ where a_j and b_j are real, then one has

$$\mathbf{w} \quad = \quad \sum_j\ a_j\,\mathbf{v}_j \ + \ b_j\,(i\,\mathbf{v}_j)$$

and therefore $\mathcal{A} \cup i \cdot \mathcal{A}$ spans W over the real numbers. To show linear independence suppose that

$$\mathbf{0} \quad = \quad \sum_j\ a_j\,\mathbf{v}_j \ + \ b_j\,(i\,\mathbf{v}_j)$$

for real coefficients a_j and b_j. If we set c_j equal to $c_j = a_j + b_j\,i$ then the diplayed equation reduces to $\Sigma_j\ c_j\,\mathbf{v}_j = \mathbf{0}$ and by the linear independence of $\mathcal{A}$ over the complex numbers this implies $c_j = 0$ for all j, which in turn implies $a_j = b_j = 0$ for all j. Therefore $\mathcal{A} \cup i \cdot \mathcal{A}$ is linearly independent over the reals, and thus we have shown that this set is a basis for W over the real numbers. This set contains $2\,k$ vectors and since the complex basis $\mathcal{A}$ for W contains k vectors we conclude that the real dimension of W is twice the complex dimension of W.■

IV.3 : Complex eigenvalues and eigenvectors

1. Let A be an $n \times n$ matrix over the complex numbers that is normal. Prove that A can be factored as the product of a unitary matrix U and a Hermitian matrix S with nonnegative real eigenvalues such that $S\,U = U\,S$. [*Hint:* Why is this true if A is diagonal? Think about what this means for 1×1 matrices. Use the Spectral Theorem to derive the general case.]

SOLUTION.

Let D be a diagonal matrix such that $d_{j,j} = z_j$. Write z_j in polar form as $r_j\,(\cos\theta_j + i\,\sin theta_j)$ and for $r_j/feq0$ and $\theta_j \in \mathbf{R}$, and set α_j equal to $\cos\theta_j + i\,\sin theta_j$. In this case we may take the matrices S_0 and U_0 to be the diagonal matrices whose (j,j) entries are r_j and α_j respectively. The matrix U_0 is unitary since it is diagonal and its diagonal entries all have absolute value equal to 1, while the matrix S_0 is real symmetric (hence Hermitian) and its eigevalues, which are merely the diagonal entries, are all nonnegative real numbers. We know that $S_0\,U_0 = U_0\,S_0$ because all diagonal matrices commute with each other.

Given an arbitrary normal matrix A, by the Spectral Theorem there is a unitary matrix P such that $P^*\,A\,P$ is a diagonal matrix D. Write $D = S_0\,U_0 = U_0\,S_0$ using the observations in the preceding paragraph, and set S and U equal to $P\,S_0\,P^*$ and $P\,U_0\,P^*$ respectively. Then the statement about the eigenvalues of S follows immediately, and since the product of unitary matrices is unitary we know that U is unitary. The commutativity of S and U follows from the sequence of equations

$$S \cdot U \;=\; \big(P\,S_0\,P^*\big)\cdot\big(P\,U_0\,P^*\big) \;=\; P\,S_0\,(P^*\,P)\,U_0\,P^* \;=\; P\,S_0\,I\,U_0\,P^* \;=\; P\,S_0\,U_0\,P^* \;=$$

$$P\,U_0\,S_0\,P^* \;=\; P\,U_0\,I\,S_0\,P^* \;=\; P\,U_0\,(P^*\,P)\,S_0\,P^* \;=\; \big(P\,U_0\,P^*\big)\cdot\big(P\,S_0\,P^*\big) \;=\; U \cdot S\,.$$

Finally, a similar string of equations

$$S\,U \;=\; \cdots\ \text{(as above)}\ \cdots \;=\; P\,S_0\,U_0\,P^* \;=\; P\,D\,P^* \;=$$

$$P\big(P^*\,A\,P\big)\,P^* \;=\; (P\,P^*)\;A\;(P\,P^*) \;=\; I\,A\,I \;=\; A$$

yields the desired factorization of A.■

IV.4 : Jordan form

Additional exercises

1. Suppose that $f(t)$ is a polynomial of degree n with rational coefficients whose highest degree term is equal to $(-1)^n\,t^n$. Show that there is an $n \times n$ matrix with rational entries whose characteristic polynomial is equal to $f(t)$. [*Hint:* Consider the matrix A whose j^{th} column is equal to the unit vector $\mathbf{e}_{j+1}$ for $j < n$ and whose last column is equal to

$$\sum_{k=0}^{n-1} (-1)^{n-1}\,c_k\,\mathbf{e}_{k+1}$$

and compute its characteristic polynomial using induction on the degree and expansion by minors.]

SOLUTION.

Following the hint, we shall proceed by induction, starting with the case $n = 1$. Let $A[c_0, \cdots, c_{n-1}]$ be the matrix given in the hint. Then the characteristic polynomial of $A[c_0]$ is easily seen to be $(-c_0 - t)$.

We shall prove by induction on n that

$$\det\bigl(A[c_0, \cdots, c_{n-1}] - t\,I\bigr) \;=\; (-1)^n\, t^n \;+\; \sum_{j=0}^{n-1} (-1)^n\, c_j\, t^j \;.$$

If we choose the coefficients c_j to be suitable rational numbers, we can realize every rational polynomial of degree n whose top degree term is $(-1)^n\, t^n$ as the characteristic polynomial of a matrix with rational coefficients.

Assuming the result in the $(n-1)\times(n-1)$ case, consider the expansion of $\det(A[c_0, \cdots, c_{n-1}] - t\,I)$ by minors along the first row. The only nonzero entries of the matrix $\det(A[c_0, \cdots, c_{n-1}] - t\,I)$ in this row are the first and last ones, and therefore the expansion by minors reduces to the following expression:

$$(-t)\cdot\det\bigl(A[c_1, \cdots, c_{n-1}] - t\,I\bigr) \;+\; (-1)^n\, c_0$$

The induction hypothesis implies that

$$\det\bigl(A[c_0, \cdots, c_{n-1}] - t\,I\bigr) \;=\; (-1)^{n-1}\, t^{n-1} \cdot \left(\sum_{j=1}^{n-1} (-1)^n\, c_j\, t^{j-1}\right)$$

and if we substitute this into the expansion by minors we obtain the formula

$$\det\bigl(A[c_0, \cdots, c_{n-1}] - t\,I\bigr) \;=\; (-1)^n\, t^n \;+\; \sum_{j=0}^{n-1} (-1)^n\, c_j\, t^j$$

that was given in the previous paragraph.■

IV.5 : Differential equations revisited

1. Let A be an $n \times n$ matrix with real entries that is orthogonal. Prove that there is a subspace $W \subset \mathbf{R}^n$ of dimension equal to 1 or 2 such that $\mathcal{L}_A$ maps vectors of W to vectors of W and also maps vectors of $W^{\perp}$ to vectors of $W^{\perp}$. [*Hint:* If A has a real eigenvalue, let W be the subspace spanned by a corresponding eigenvector. If A does not have a real eigenvalue, then it does have a complex eigenvalue λ with a corresponding eigenvector that can be written in the form $\mathbf{x} + i\,\mathbf{y}$ where both $\mathbf{x}$ and $\mathbf{y}$ have real coordinates. We also know that $\overline{\lambda}$ is an eigenvalue with eigenvector $\mathbf{x} - i\,\mathbf{y}$. Using these facts, show that $\mathcal{L}_A$ sends the real subspace spanned by $\mathbf{x}$ and $\mathbf{y}$ into itself.]

SOLUTION.

Let $\mathbf{x} + i\,\mathbf{y}$ be an eigenvector for A over the complex numbers with associated eigenvalue $a + b\,i$. Then we have

$$A\mathbf{x} \;+\; i\,A\mathbf{y} \;=\; A\,(\mathbf{x} \;+\; i\,\mathbf{y}) \;=\; (a + b\,i)\,(\mathbf{x} \;+\; i\,\mathbf{y}) \;=\; \bigl(a\,\mathbf{x} - b\,\mathbf{y}\bigr) \;+\; i\cdot\bigl(b\,\mathbf{x} + a\,\mathbf{y}\bigr)\;.$$

Equating the real and imaginary parts of the first and last expressions, we see that $A\,\mathbf{x} = a\,\mathbf{x} - b\,\mathbf{y}$ and $A\,\mathbf{y} = b\,\mathbf{x} + a\,\mathbf{y}$. If W denotes the span of $\mathbf{x}$ and $\mathbf{y}$ then it follows that $\mathcal{L}_A$ maps W into itself.

In fact, we claim that $\mathcal{L}_A$ maps W *onto* itself. This amounts to showing that the image of a basis for W is linearly independent. But suppose that the vectors $\mathbf{w}_j$ form a basis for W and that we can express $\mathbf{0}$ as a linear combination $\sum_j \; c_j\, A\,\mathbf{w}_j$. We then have

$$\mathbf{0} \quad = \quad \sum_j \; c_j\, A\,\mathbf{w}_j \quad = \quad A\left(\sum_j \; c_j\,,\mathbf{w}_j\right)$$

and since A is invertible this implies $\sum_j \; c_j\,\mathbf{w}_j = \mathbf{0}$. By linear independence we must then have $c_j = 0$ for all j.

Next, to show that $\mathcal{L}_A$ maps $W^\perp$ into itself, we need to show that if $\langle \mathbf{x},\, \mathbf{w}\rangle = 0$ for all $\mathbf{w} \in W$ then $\langle A\,\mathbf{x},\, \mathbf{w}\rangle = 0$ for all $\mathbf{w} \in W$. By the preceding paragraph we may write $\mathbf{w} = A\,\mathbf{z}$ for some $\mathbf{z} \in W$. Since A preserves inner products we then have

$$\langle A\,\mathbf{x},\, \mathbf{w}\rangle \quad = \quad \langle A\,\mathbf{x},\, A\,\mathbf{z}\rangle \quad = \quad \langle \mathbf{x},\, \mathbf{z}\rangle$$

and the right hand side is equal to zero because $\mathbf{z} \in W$ and we assumed that $\mathbf{x}$ was orthogonal to every vector in W.∎

2.** Using mathematical induction and the result from the previous exercise, prove that if A is an orthogonal matrix then there is an orthogonal matrix P such that ${}^{\mathbf{T}}P\,A\,P$ is a block sum of orthogonal matrices where each summand is either 2×2 or 1×1.

SOLUTION.

We shall formulate the exercise in terms of linear transformations. Suppose that V is a finite dimensional real inner product space and T is an orthogonal linear transformation from V to itself. We need to show that there is an orthonormal basis of V consisting of vectors $\mathbf{e}_1, \cdots, \mathbf{e}_p$ and $\mathbf{f}_1, \cdots, \mathbf{f}_r$ where $r \leq p$ and T maps the subspaces W_j spanned by $\mathbf{e}_j$ and $\mathbf{f}_j$ to themselves if $j \leq r$ and the subspaces W_j spanned by $\mathbf{e}_j$ to themselves if $j > r$.

As in several of the proofs in Section IV.3, we proceed by induction on the dimension of V. If this dimension is 1 or 2, then the conclusion is immediate; assume we know the result for all orthogonal linear transformations on reall inner product spaces of dimension strictly less than $\dim V$.

The preceding exercise shows that there is a subspace V_0 of dimension 1 or 2 such that T maps both V_0 and its orthogonal complement $V_0^\perp$ into themselves. As in similar cases from the notes, let S be the linear transformation from $V_0^\perp$ to itself that is defined by $S(\underline{\mathrm{y}}) = T(\mathbf{y})$. Then the induction hypothesis applies to S on V_0. There are two cases depending upon whether $\dim V_0 = 1$ or 2.

If $\dim V_0 = 1$ then $\dim V_0^\perp = \dim V - 1$ and there is an orthonormal basis for $V_0^\perp$ consisting of vectors $\mathbf{e}_1, \cdots, \mathbf{e}_{p-1}$ and $\mathbf{f}_1, \cdots, \mathbf{f}_r$ where $r \leq p-1$ and T maps the subspaces W_j spanned by $\mathbf{e}_j$ and $\mathbf{f}_j$ to themselves if $j \leq r$ and the subspaces W_j spanned by $\mathbf{e}_j$ to themselves if $j > r$. If we let $\mathbf{e}_p$ be a nonzero unit vector in V_0, then adding it to the given orthonormal basis for $V_0^\perp$ will yield an orthonormal basis for V that has the desired properties.

On the other hand, if $\dim V_0 = 2$ then $\dim V_0^\perp = \dim V - 2$ and there is an orthonormal basis for $V_0^\perp$ consisting of vectors $\mathbf{e}_2, \cdots, \mathbf{e}_p$ and $\mathbf{f}_2, \cdots, \mathbf{f}_r$ where $r \leq p$ and T maps the subspaces W_j spanned by $\mathbf{e}_j$ and $\mathbf{f}_j$ to themselves if $j \leq r$ and the subspaces W_j spanned by $\mathbf{e}_j$ to themselves if $j > r$. If we take an orthonormal basis for V_0 consisting of vectors $\mathbf{e}_1$ and $\mathbf{f}_1$, then adding them to the given orthonormal basis for $V_0^\perp$ will yield an orthonormal basis for V that has the desired properties. This completes the proof of the inductive step in the argument.∎

3. Let A be a 2×2 matrix with real entries.

(a) If the characteristic polynomial is a product of linear factors over the real numbers, explain why A is similar to a matrix in Jordan form with real entries down the diagonal.

SOLUTION.

There are two cases depending upon whether the polynomial has two distinct real roots or a single root with algebraic multiplicity 2. In the first case we know that there is a basis of eigenvectors, and if P is the matrix whose columns are the eigenvectors, then $P^{-1}AP$ is a diagonal matrix whose entries are the two real eigenvalues. The conclusion of the exercise follows because diagonal matrices are a special case of Jordan forms. Also, if the characteristic polynomial has a double root and a basis of eigenvectors, the same argument goes through, so we are left with the case where there is a double root but A is not diagaonalizable. Write the characteristic polynomial as $(\lambda - t)^2$. In this case the matrix $N = A - \lambda I$ has real entries and since $N^2 = 0$ by the Cayley-Hamilton Theorem, the argument at the beginning of this section (in the notes) shows that there is an ordered basis consisting of vectors $\mathbf{x}$ and $\mathbf{y}$ with real coordinates such that $N(\mathbf{y}) = \mathbf{x}$ and $N(\mathbf{x}) = \mathbf{0}$. It follows immediately that $A\mathbf{x} = \lambda\mathbf{x}$ and $A\mathbf{y} = \mathbf{x} + \lambda\mathbf{y}$, and therefore the matrix of A with respect to the ordered basis determined by $\mathbf{x}$ and $\mathbf{y}$ is equal to

$$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} . \blacksquare$$

(b) If the characteristic polynomial is not a product of linear factors over the real numbers and $\det A > 0$, show that A is similar over the real numbers to a matrix of the form

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

where the complex eigenvalues of A are $a \pm bi$.

SOLUTION.

In this case there are two nonreal eigenvalues that come in complex conjugate pairs, and we shall denote them by $p \pm qi$. Let $\mathbf{v}$ and $\mathbf{w}$ be vectors in $\mathbf{R}^2$ such that $\mathbf{v} \pm i\,\mathbf{w}$ is a complex eigenvector for $p \pm qi$. As in Exercise 1, it follows that $A\mathbf{v} = p\mathbf{v} - q\mathbf{w}$ and $A\mathbf{w} = q\mathbf{v} + p\mathbf{w}$. If $\mathbf{v}$ and $\mathbf{w}$ are linearly independent, then the conclusion will hold for the ordered basis given by $\mathbf{v}$ and $\mathbf{w}$ if we take $a = p$ and $b = -q$, so the proof reduces to showing that the two vectors in question are linearly independent.

Since we have a real matrix with nonreal eigenvalues and $\mathbf{v} \pm i\,\mathbf{w}$ is the complex eigenvector for $p \pm qi$, then $p + qi \neq p - qi$ implies that $\mathbf{v} + i\,\mathbf{w}$ and $\mathbf{v} - i\,\mathbf{w}$ are linearly independent. We shall use this to prove that $\mathbf{v}$ and $\mathbf{w}$ are linearly independent. Suppose that we have $c\mathbf{v} + d\mathbf{w} = \mathbf{0}$ for some scalars c and d. Rewriting $\mathbf{v}$ and $\mathbf{w}$ as linear combinations of the vectors $\mathbf{u}_\pm = \mathbf{v} \pm i\,\mathbf{w}$ by the formulas

$$\mathbf{v} = \tfrac{1}{2}(\mathbf{u}_+ + \mathbf{u}_-) \quad , \quad \mathbf{w} = \tfrac{1}{2i}(\mathbf{u}_+ - \mathbf{u}_-)$$

and using the linear independence of $\mathbf{u}_\pm$ we see that

$$0 = \frac{1}{2}c + \frac{1}{2i}d = \frac{1}{2}c - \frac{1}{2i}d .$$

The only solution for this system of equations is $c = d = 0$, and therefore the vectors $\mathbf{v}$ and $\mathbf{w}$ must be linearly independent as required.$\blacksquare$